【07】第三章 | 进阶下载，从半中心化到完全去中心化（二）

BT，种子，磁力链接，magnet:?xt=，相信每一个经常下载文件的人看到这些名词都会眼前一亮。没错，这些词都属于我们在使用 BT（BitTorrent，种子）下载时频繁碰到的单词，基本上已经达到了习以为常的程度。也许你心中认为对它们已经相当地了解，那么当一个下载新手虔诚地向你这个「高玩」提问这些名词时你是否又能说出来呢?

如果你有一种「自己好像能说清楚又不知道怎么说」的感觉，在看完这节之后想必会豁然开朗，不仅能更好地理解我们之后教到的小技巧，还有机会在小白面前炫耀一番。

从种子文件理解 BitTorrent 协议

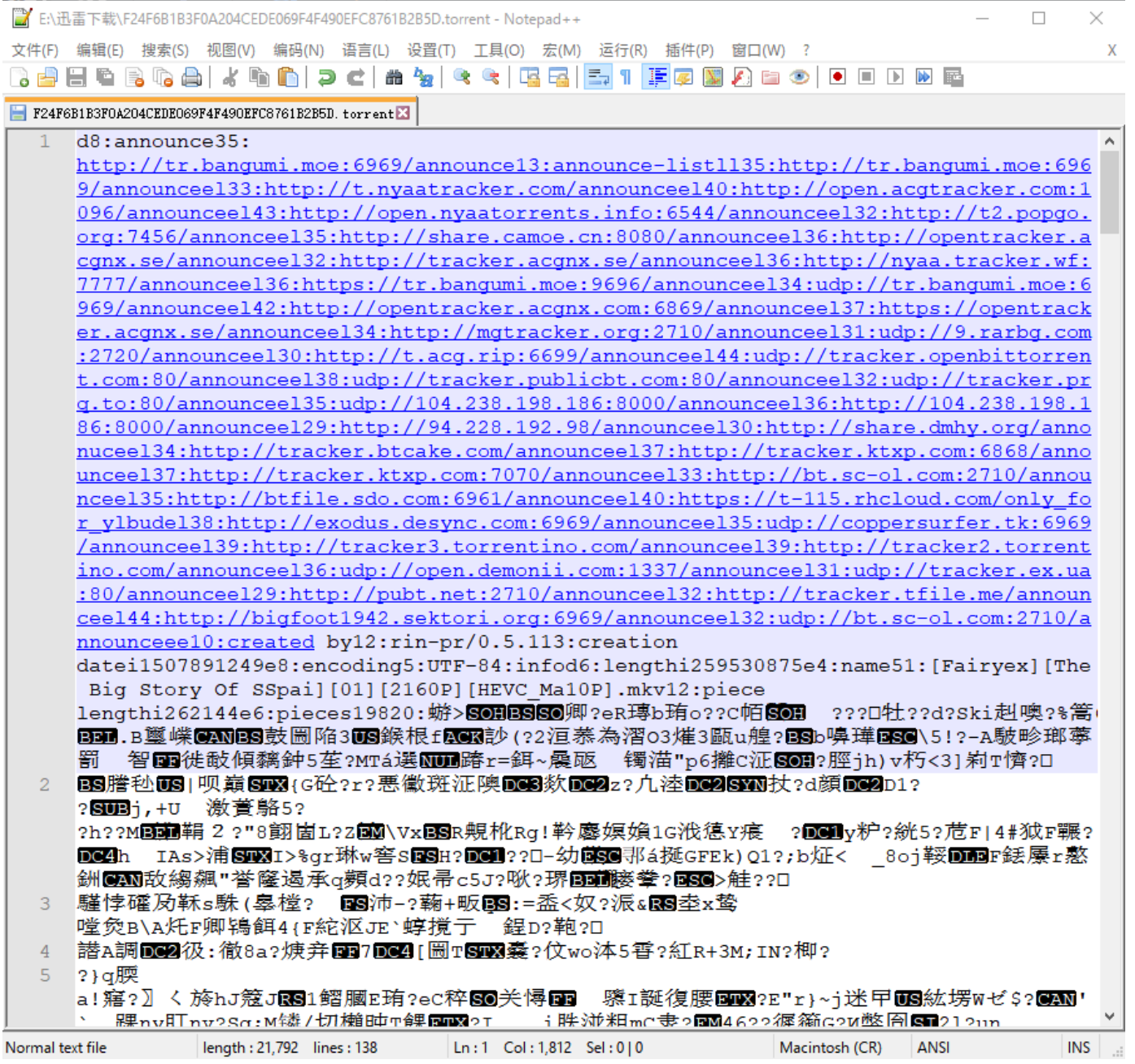
说到 BT 就不得不说扩展名为 .torrent 的种子文件（英语：Torrent file），因为必须有种子文件才可以下载到对应的 BT

资源。那平时老听到有人说留种的时候，到底留的是什么呢?

留的是 Tracker 信息和文件信息。

为了方便理解，我们以一部虚构电影《少数派大电影》的种子 [Fairyex][The Big Story Of SSpai][01][2160P][HEVC_Ma10P][MKV].torrent 为例。

当用记事本或者其他文本浏览器打开这个种子文件，看到的是一大串类似于链接的字符串，后面跟着的是谁都看不懂的乱码（大家也可以自己试一下打开电脑上的种子文件）：



Tracker 信息

蓝色的链接即为这个种子自带的下载中需要用到的 Tracker 服务器的地址和针对 Tracker 服务器的设置，Tracker 服务器的作用就是提供给下载者已经下载或者正在下载少数派大电影的其他下载者的地址信息，方便下载者与其他下载者连接。

种子内置了 Tracker 的好处就在于如果大家都是用这个种子

下载的少数派大电影，后来下载的人可以直接通过这几个 Tracker 和正在下载或拥有这个文件的人连接。

文件信息

在 Track 服务器之后我们还能看到一段不是乱码的字符，这里和后面的乱码一起组成了需要下载的文件信息。

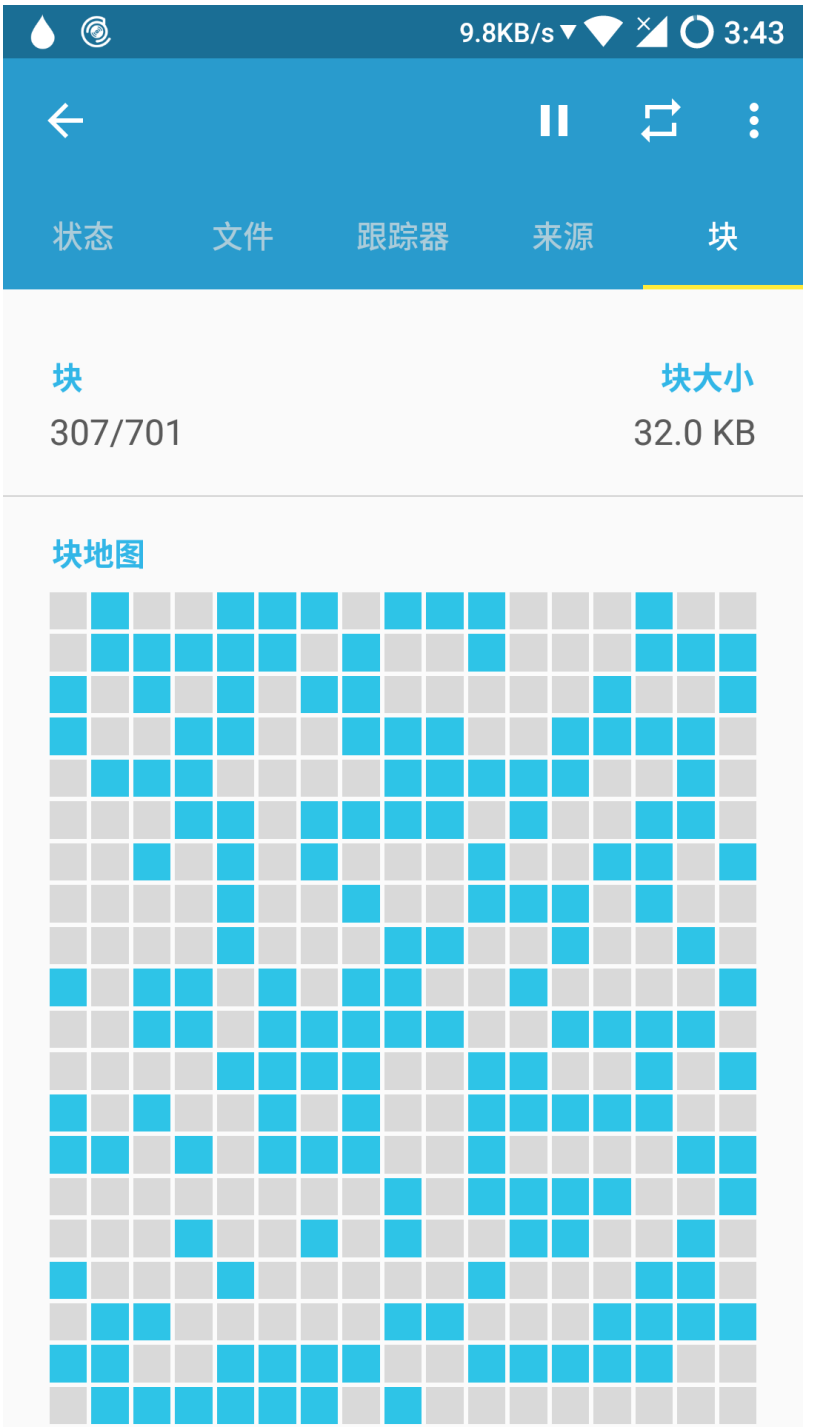
先简单说一下常见的文件基础信息，信息以冒号分割：

- creation date: 种子创建日期，是从1970年1月1日00:00:00到现在的秒数。
- encoding: 种子文本编码，一般都是 UTF-8。
- name: 显而易见 [The Big Story Of SSpai][01][2160P][HEVC_Ma10P].mkv 是我们要下载的文件名。
- Pieces: 文件被分成了多少块（可以看到这里分为了 32240 块）。
- piece length: 每块文件的大小。
- files 与 path: 如果种子文件中有多个文件或文件夹，这两个关键字就是用来展示它们的。

小知识：种子文件里面也可以包含 eD2k 信息。

剩下的乱码即是 BT 生成的每一块切好的文件块的指纹，在下载少数派大电影的时候每下载一块（这里是 1/32240）下

载器会对比种子里的指纹和下载数据是不是匹配的，这样子可以确保下载时数据的准确性。即使是出错，需要重下的也只是出错的那小小的一部分（1/32240）。



一些下载 App 可以看到文件块的情况

BT 下载流程

好了，了解清楚种子的结构之后再去理解 BT 下载就如探囊取物一般轻松自如啦。还是以少数派大电影为例：

1. 在一个支持解析 BT 协议的下载客户端（如迅雷）中打开 [Fairyex][The Big Story Of SSpai][01][2160P][HEVC_Ma10P][MKV].torrent，迅雷开始读取文件信息并显示下载提示（所以平时我们打开种子的时候文件名列表和大小都不是联网获取的哦）。

2. 迅雷联系种子内置的 Tracker 服务器们，把少数派大电影的文件指纹发过去，告诉 Tracker 服务器们我要下载这个文件啦，服务器收到后首先记录下你（以 IP 为身份）正在下载这个文件（记录种子信息），然后把它知道的正在下载同一部电影的人的 IP 地址返回给迅雷。

当然如果其他正在下载少数派大电影的人请求 Tracker 服务器的话它也会把你的 IP 发过去哦，你长时间不和 Tracker 服务器联系的话它会自动把你从正在下载的人中除名。那如果一个正在下载或已经下载下完少数派大电影的人也没有怎么办？没办法，迅雷只能一直和服务器联系，直到有缘人出现。

迅雷收到这一大堆 IP 地址之后，会过滤掉其中连接不上的，接着开始和连接上的人互相商量，内容大概就是（实际下载中数字是随机的）：

迅雷：我要下载少数派大电影，可是现在我连一个文件块都没有... 嘤嘤嘤。

下载者 A：我有第 1,3,4,5,7,10...32239 块，我给你1,3 块吧。

迅雷：我有少数派大电影的第 1,3 个文件块啦，我要第 2,4,5,6...32240 块哦。

下载者 B：我有第 2 块，我给你吧，我要第 3,6 块。

迅雷：好的，不过没有第 6 块只能给你第 3 块。

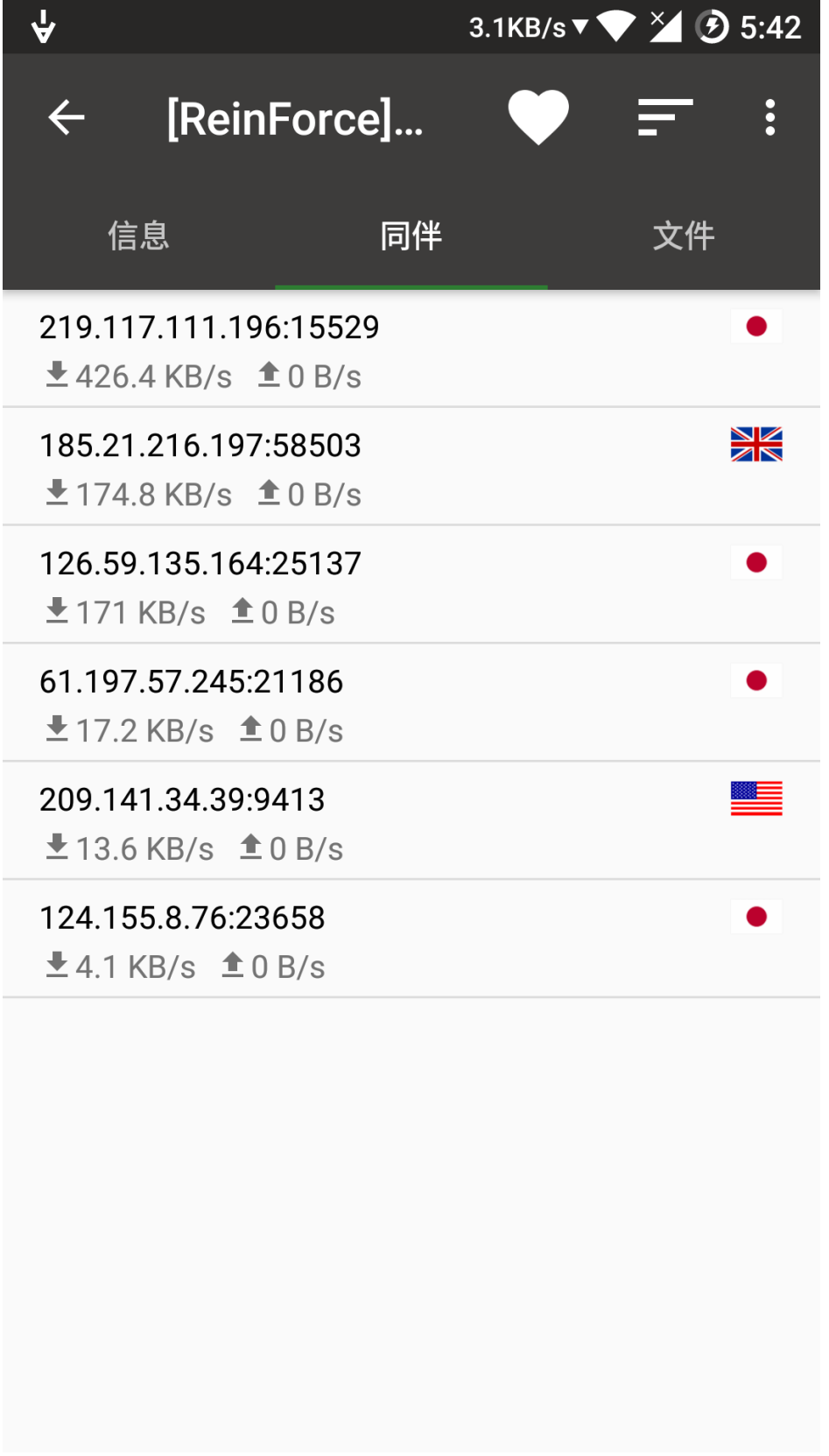
下载者 C：我有第 5,8 块，我给你吧，你把你的第 10,32239 块给我吧。

迅雷：好的。

...

下载者 N：我有第 5601,7979...26654 块，我给你吧，你把你的第 1,4,5... 32240 块给我吧。

迅雷：虽然我已经下完了，还是发给你吧。



从 Aria2 可以看到好几个国家的下载者正在和我交换数据

就这样，在这样你帮我我帮你的「愉快」氛围下大家下载少数派大电影的速度都轻松飙满，不一会儿就全下完啦。不过

即使你已经下载完成，这种对话也会一直持续，直到你停止上传或退出客户端为止。有一个重点是这些文件数据交换并不是按照某一个特定的间隔或顺序发生的，每个用户（peer）和其他用户之间都会发生这些对话，视觉化之后有点像我们小时候做的「N 点之间有多少连线」的数学题。

资源的诞生 —— 做种

前面那么多都是解释用户是怎么下载的，那么提供下载的文件是怎么来的呢？这就需要第一个拥有这个文件的人制作种子（简称做种）。还是少数派大电影，现在它正在上映，有个「电影小偷」拿着迷你 DV 进去电影院最后一排偷偷地把整部电影拍了下来并且压制成了传说中的「高清 TS 中字」版，回到家之后这个「电影小偷」打开了可以制作种子的 BitTorrent 客户端（例如 Bitcomet，中文名比特彗星）选择压制好的电影生成种子文件。

接下来「电影小偷」挑了一个人多的资源网站发了个帖子并把种子附了上去，很快就有用户开始下载。在耐心等到第一批用户下载完成之后，「电影小偷」直接把自己电脑上的电影文件和种子文件都删掉，这样谁都找不到他，深藏功与名。

虽然我们一直在举如何下载盗版的例子，不过 BT 下载可是为了正经用途而生的，它也经常被用来做正经的用途（不止下载）。例如游戏和系统更新这种短时间内有大量用户下载的大文件，通常大小都是以 GB 为单位，如果全部使用传统的 HTTP 下载，那把公司买破产了都不一定能买到那么多带

宽；即使带宽足够，等更新高峰过后这些带宽又会被闲置，造成很大的资源浪费。

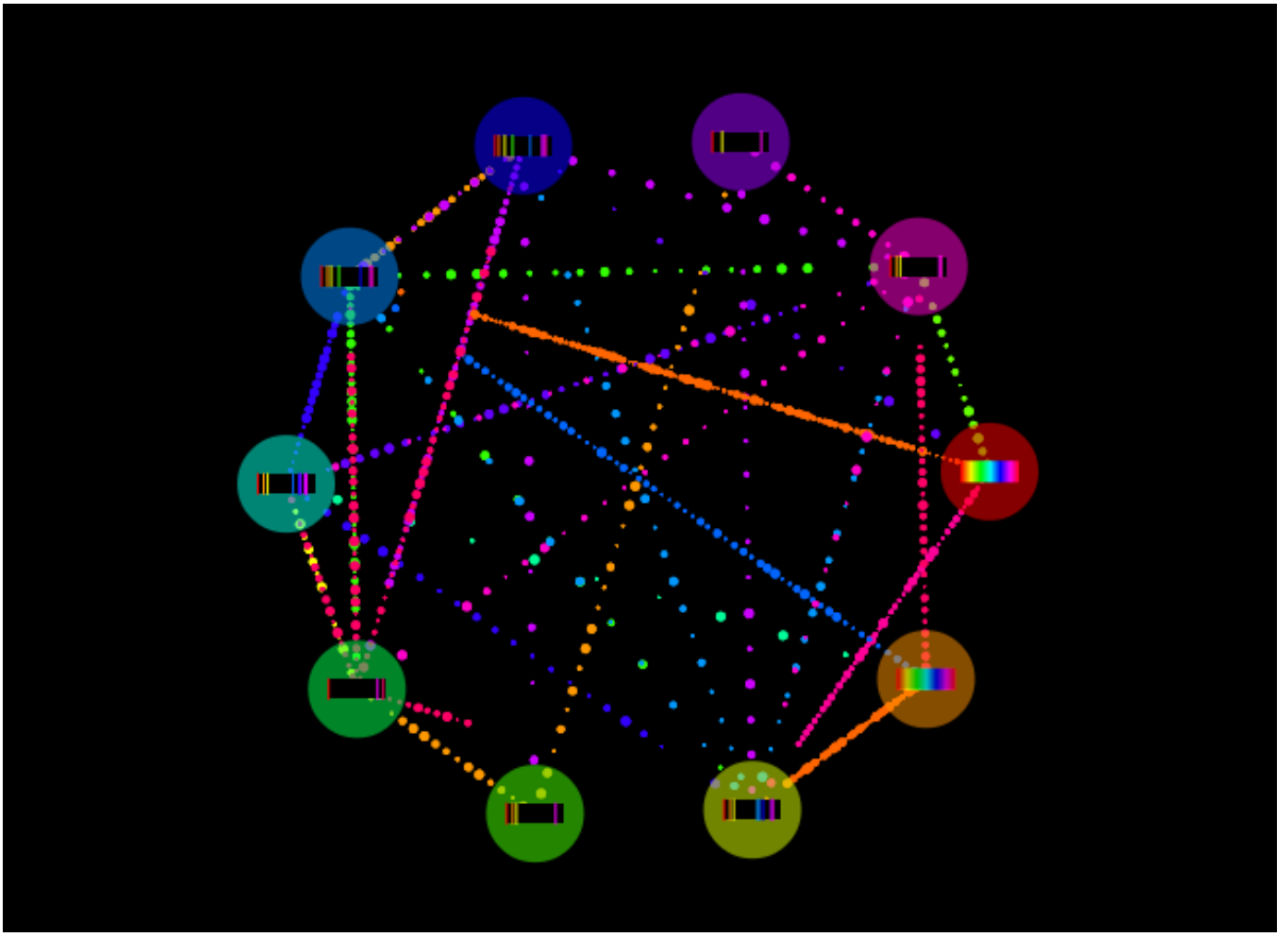
现代很多的游戏和系统都会内置一个 BT 下载器用来下载大文件，数据只在用户之间传播，服务器只需要负责做种和 Tracker，大大节省了成本和压力。除了下载，你一定想不到许多免费甚至收费的「爱国上网」APP 也是用了 BT 的原理，所以这些「良心 APP」的实际运营成本并没有你想象的那么多。至于其他使用 BT 下载的领域，就留给你自己探索咯。

以魔兽世界那时候的玩家量和网络条件，不用 BT 暴雪需要花多少钱？

上帝视角观察下载流程

这里专门为「太长不看」党和「还是一脸懵逼」党准备的小节，以下这个地址提供了一个直观的 BT 下载原理演示，只需要动动小手轻松理解 BT 下载：

http://mg8.org/processing/bt.html



附操作翻译：按键盘上的 s 键添加一个下载完成的「种子用户」，按 p 键添加一个嗷嗷待哺的「0% 用户」，按 r 键随机删掉一个节点（需要几秒等待模拟数据传完才会删除）。

BT 的「救星」—— 磁力链接

在前面的流程里面，大家有没有看出什么问题呢?

没错，就是 Tracker 服务器。它在 BT 下载的过程中起着不可或缺的作用，没了它我们就找不到和我们一起下载的小伙伴，相当于 BT 下载的「中心」，这个时候的 BT 下载和早期 eD2K 一样只能说是「半中心化」。由于 BT 下载经常被人用来传播无版权和情色限制级资源，Tracker服务器很容易被迫关闭或被墙。

举个很多人都知道的例子 —— 全球最大的 BT 种子服务器海盗湾，它同时也是一个巨型 Tracker，无数合法或不合法的资源都能够在海盗湾找到。正所谓树大招风，它被所在国瑞典的警察反复查水表，靠着转移服务器和将种子转换为磁力链接坚强地活了下来。而国外大部分 Tracker 就没那么好运了，关闭的关闭停止解析的停止解析。

在中国我们也可能需要担心 Tracker 的问题，一方面国外的 Tracker 都被封得差不多了，只不过法律的不完善和自我阉割让迅雷这个超巨型 Tracker 和其他一些小 Tracker 活下来。另一方面这些干干净净的 Tracker 可以下载的文件的确有限，一旦资源过气或者迅雷自我阉割对于国内用户来讲就几乎等同于不能下载。

Tracker 服务器自带的另一个严重缺点是所有文件分享者都会被它追踪，特别是在中国。关于追踪和监控的详情会在关于下载的安全性问题章节中详谈。

还记得在第二节中小明的例子吗，这个例子里面是没有「中心」的完全去中心化，一些聪明人就通过基于 KAD（Kademlia）网络的DHT（Distributed hash table，分布式哈希表）网络技术实现了这个例子，磁力链接就是基于 DHT 诞生的。

磁力链接（Magnet URI scheme）的出现让 Tracker 不再是 BT 下载中的必需品，甚至不再需要 Tracker。很好地解决了之前 BT 下载的缺陷，我们现在就来解释一下磁力链接是如何做到只需要短短一个链接就能实现之前一个种子文件才能实现的功能。

磁力链接的结构

magnet:?
xt=urn:sha1:YNCNHJKIWBTRNJIV4YUAE52SJUQFGY4A

注：虚构链接请勿下载

除了以 magnet:? 开头，我们经常还可以在连接前面看到一个小磁铁的图标，其中 xt 为必须项，后面跟着是根据文件或文件夹内容生成的身份 —— 数字指纹。其他常见选填项还有：

- dn：文件名

- xl：文件大小
- xs：P2P链接
- Kt：用于搜索的关键字
- Mt：文件列表
- Tr：BT下载的 Tracker 地址

磁力链接的原理

DHT 的原理可以用一句话来概括：把所有的用户都变成 Tracker。和传统的 Tracker 一样，每一个支持磁力链接的下载客户端都会在本地以一个名为 DHT.dat 的文件动态地储存和更新一小段下载者和下载文件的信息，可以简单理解为如果正在联网的下载客户端足够多，那么 Tracker 信息合并起来就相当于一个巨大的 Tracker 服务器。

接上上一个故事，少数派发现自己的大电影盗版在网上已经铺天盖地，于是赶紧报警。警方利用搜查证提取了最多人下载的几个 Tracker 上的记录，过滤出了前几个提供数据的 IP，又通过运营商查询 IP 访问记录确定了最早提供数据的 IP。恰好这个倒霉的电影小偷是在国内偷拍做种的，接下来就是和谐资源，抓人没收，定罪坐牢执法三连，少数派大电影的盗版也暂时销声匿迹。

同样身处中国的电影小偷 B 也偷拍了少数派大电影，不过他非常地聪明，做种的时候生成的是磁力链接。有了磁力链接

（如 magnet:?xt=urn:sha1:YNCNHJKIWBTRNJIV4YUAE52SJUQFGY4A）之后电影小偷 B 并不会直接把链接发到公开论坛上，而是这样子做：

1. 先找到一个到处求大电影资源的用户 A，利用 Telegram 等私密通信的方式把磁力链接发给他。

2. 用户 A 在下载客户端（以迅雷为例）打开磁力链接，迅雷查询自己 DHT.dat 文件中的客户端向他们询问有没有文件指纹为 YNCNHJKIWBTRNJIV4YUAE52SJUQFGY4A 的文件呀，结果当然是没有啦（假设没有其他做种者并且和谐得非常彻底）。然后这些客户端会和迅雷回复：你别急，我再帮你问问，接着这些客户端也会查询自己 DHT.dat 文件中的客户端并向他们请求这个文件，这样子一直到发现电影小偷 B 做种的客户端为止。（实际时间并不会很长，几秒钟左右）

3. 用户 A 的迅雷接收到原路返回的联系信息之后就开始连接电影小偷 B 的客户端进行下载，这时候电影小偷 B 如法炮制地把链接发给用户 C，D，E...（实际上不一定会一个个发，有专门发匿名信息的论坛和网站，一些用户会主动盯着这些论坛和网站）。

4. 在有一定基数的用户下载完少数派大电影后电影小偷 B 关闭客户端轻松隐身，再接下来的用户的客户端通过 DHT 搜索到的就是其他用户了，这里即是传统的 BT 下

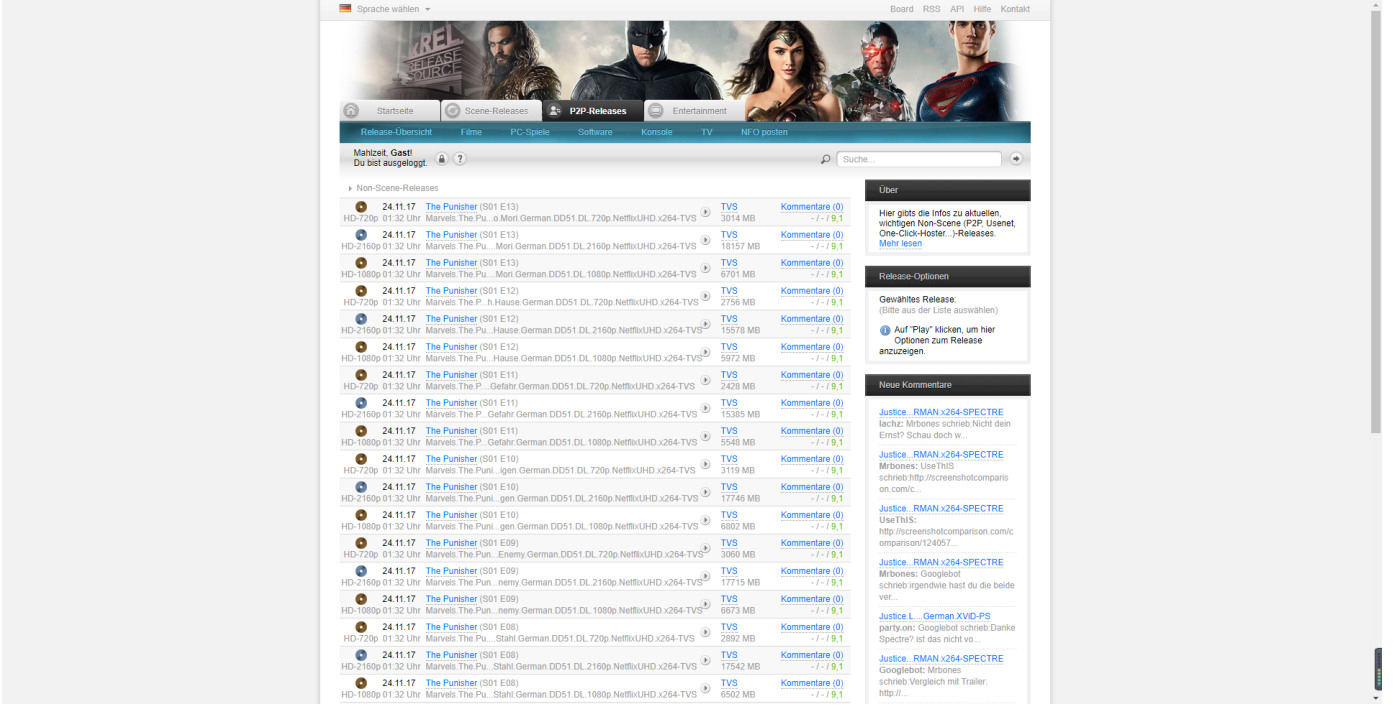
载。

通过磁力链接电影小偷 B 暴露了什么呢？一个文件指纹 YNCNHJKIWBTRNJIV4YUAE52SJUQFGY4A，向最初的几个用户暴露了自己的地址，就这么多信息。因为没有中心服务器，不像传统的做种者一样会被 Tracker 记录上传信息，所有的通讯联系都是通过客户端间进行的，由于没办法对全中国支持磁力链接的下载客户端筛选连接记录，警方想要通过追查上传源来找到电影小偷 B 几乎不可能。

而除了电影小偷 B 以外所有的下载用户都可以随意分享少数派大电影的磁力链接，因为不是上传者，传播的也只是一串字符而已（像是经常听到老司机不怕发「车牌号」也是这个原因），资源就这样被传播开来了。除了从搜索引擎或者约谈站长和谐磁力链接，少数派面对磁力链接也只能干瞪眼。

当然磁力链接也不是绝对地安全，更多有关磁力链接安全性的问题会关于下载的安全性问题章节中详谈。

轻巧，相对安全和去中心化的优点让越来越多的网站删除种子开始全面使用磁力链接，越来越多的破解软件和游戏破解组，还有各种不想公开身份的汉化组也都开始从 Topsite（简单理解: 私人限制进入的 FTP 服务器）转为磁力链接分享资源，在今后这一趋势还会持续增长。



许多知名的破解组（比如大名鼎鼎的 CPY，不了解的同学请谷歌 CPY 破解）都是在这里发布资源的

新篇章

至此本篇科普与教程所涉及的所有下载方式原理已经科普结束啦，首先给能够坚持阅读并且成功转化为自己见识的各位鼓掌。掌握了原理，再看我们接下来讲的内容那是非常简单直接，充满趣味的。

下面一节内容我会说一下未来有机会代替 BT 下载的各种有前途但是目前还不为大众所知的下载方式，不会带有太多专有名词和巨长无比的解释，权当给大家放松一下。